Introduction to Statistical Methods in Political Science

Lecture 4: Introduction to Probability

Ignacio Urbina
Ph.D. Candidate in Political Science

Random Processes, Outcomes, and Events

Why Study Probability?

- We want to understand how likely an event is.
- Compare events to say which is more likely.
- When we say an event is more likely, we expect it to happen more than less likely events.
- Likelihood and probability reflect our belief about the chances of events occurring.
- Examples:
  - We can say very confidently that the sun will rise tomorrow or "*there is a very high probability the sun will rise tomorrow*."
  - We can say very confidently that in the middle of summer, it won't snow or "*There is a low probability of snow in the middle of summer*."

A Solid Understanding of Probability

- Provides a framework to deal with uncertainty and randomness.
- Our expectations about likely events influence our actions and planning.
- Examples:
  - Planning an outdoor event based on the weather forecast.
  - Buy a stock based on its price forecast.
  - Settle on a court case based on the probability of losing the lawsuit.
- We study probabilities of **random processes**.

What is a Random Process?

Random Process (*Def.*)

A random process is one where the outcome is uncertain before it happens.

- The outcome is drawn from a set of all possible outcomes.
- Examples:
  - Weather: We cannot predict with certainty if it will rain tomorrow.
  - Throwing a die: The result of a die throw is unknown until it happens.
  - Elections: The winning candidate is uncertain until votes are counted.

What is a Random Process?

- What do we mean by "process"?
- **Process:**
  - A process is a mapping from an initial state to a later state, i.e., a sequence of connected steps.
  - Example: Applying for a scholarship (learning about it, writing the application, submitting it, getting the result).
  - Initial conditions lead to a new condition through a series of steps.
- **Random Process:** A random process involves initial conditions leading to an uncertain outcome.

Example: Election as a Random Process

- Initial conditions: Current political climate, economic conditions, media coverage.
- These factors influence voter preferences in complex ways.
- Outcome: Election result (winning candidate).
- Unpredictable events during the campaign can change voter behavior.
- New information, debates, and events can shift the final outcome.

Understanding Outcomes

Outcome (*Def.*)

An outcome is one realization of a random process.

- Examples:
  - Rain or no rain tomorrow.
  - Rolling a one on a die.
  - "Candidate A" winning the election.

What is an Event?

Event (*Def.*)

An event is a collection of outcomes from a random process.

- Event groups outcomes together so that we can study the probability of broader conditions and scenarios rather than just isolated cases.
- Events can represent conditions that occur in various contexts and times (e.g., different days, multiple experiments).

Event vs. Outcome

- An outcome is a single realization of a random process.
- An event is a set of possible outcomes.
  - **Rolling a die**:
    - *Outcome:* Getting a 6 when throwing a die.
    - *Event:* Getting an even number when rolling a die (2, 4, or 6).
  - **Electoral Results**:
    - *Outcome:* Candidate A wins narrowly.
    - *Event:* Candidate A wins (including both "Wins by a landslide" and "Wins narrowly").

Example: Forecast Scenarios 2024 Election

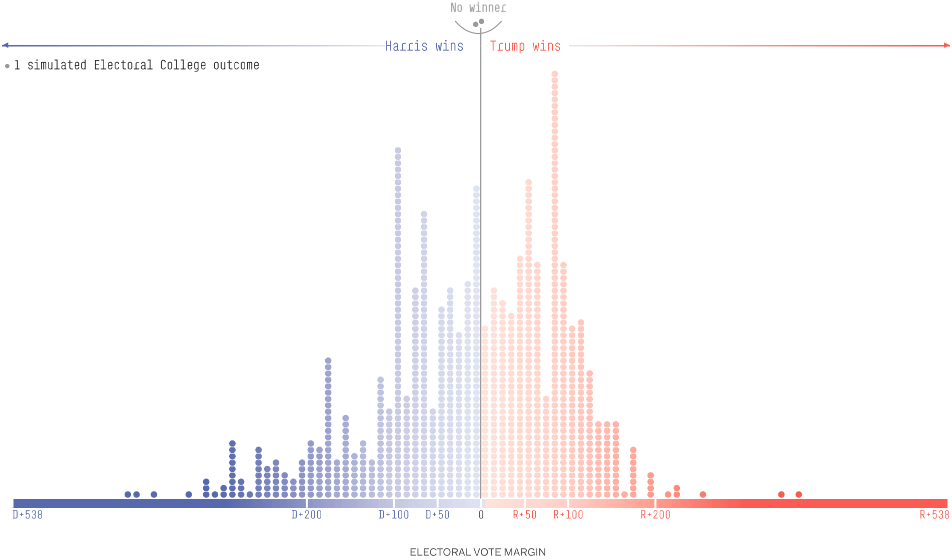


Figure: Source: 538

Disjoint Events

- **Disjoint (Mutually Exclusive) Events:** Two events are disjoint if they do not share any outcomes.
  - Example: Rolling a die
    - Event A: Rolling an even number $\{2, 4, 6\}$
    - Event B: Rolling an odd number $\{1, 3, 5\}$
    - These events are disjoint because they do not share any outcomes.
  - Example: Election (First-past-the-post voting system)
    - Event A: Any candidate from Party X wins $\{X_1, X_2, \cdots\}$
    - Event B: Any candidate from Party Y wins $\{Y_1, Y_2, Y_3, \cdots\}$
    - These events are disjoint because they do not share any outcomes.

Join Events (Not Disjoint)

- **Joint Events (can occur together):** Two events are joint if they share at least one outcome.
  - Example: Weather
    - Event A: It rains tomorrow $\{Rain\}$
    - Event B: It is windy tomorrow $\{Windy\}$
    - These events are joint since the outcome is $\{Rain, Windy\}$ can happen.
  - Example: Rolling two dice
    - Event A: Rolling a 2 on the first die $\{2\}$
    - Event B: Rolling a 4 on the second die $\{4\}$
    - These events are joint because the combined outcome $\{2, 4\}$ can occur.

Set Theory Basics

Set Theory - Basic Definitions

Definition: A set is a collection of well-defined, unordered objects called elements or members.
Explicit Set Definition: A (finite) set can be defined by explicitly specifying all of its elements between curly braces, known as set braces $\{\}$.
For example,

- $A = \{1, 2, 3, 4, 5, 6\}$
- $B = \{a, e, i, o, u\}$
- $C = \{US, UK, FRANCE, CANADA, CHINA, ...\}$
- $D = \{h, t\}$ (outcomes of a coin toss: heads, tails)
- $S = \{\{h, h\}, \{h, t\}, \{t, h\}, \{t, t\}\}$ (outcomes of two coin tosses)

Set Theory - Basic Definitions

Review:

- The null set (or empty set) is denoted by $\emptyset$.
- The union of sets $C \cup D$ includes elements in C, D, or both.
- The intersection of sets $A \cap B$ includes elements common to both A and B.
- The complement of a set $\neg D$ or D^c includes all elements not in D.
- Mutually exclusive (or disjoint) events are sets with no common elements.
- A series of exhaustive events cover all possible outcomes in the sample space.
- The universe set, U, is the set that contains all possible elements.

Set Theory Basics

Define the sets

- $A = \{\text{Banana, Apple, Orange, Watermelon}\}$,
- $B = \{\text{Orange, Plum, Grapes, Apple}\}$.
- Suppose the universal set is $U = \{\text{Banana, Apple, Orange, Watermelon, Plum, Grapes, Lemon}\}$.

Then,

- $A \cup B = \{\text{Banana, Apple, Orange, Watermelon, Plum, Grapes}\}$
- $A \cap B = \{\text{Apple, Orange}\}$
- Complement of $A = A^c = U - A = \{\text{Plum, Grapes, Lemon}\}$
- $A \cup A^c = U$.

Probability Definitions, Axioms, and Probability Distribution

Definitions of Probability

- **Frequentist Definition**

Definition

The probability of an event is the proportion of times the event would occur if we observed the random process an infinite number of times.

- Probability is defined as the long-run frequency of an event occurring.
- Example: Flipping a coin many times and observing the proportion of heads.

Definitions of Probability

- **Classical Definition**

Definition

The probability of an event is the number of ways it can happen divided by the total number of possible outcomes, assuming all outcomes are equally likely.

- Example: Rolling a fair six-sided die, the probability of getting a number lower than 3 (so, either 1 or 2) is $\frac{1+1}{6} = \frac{2}{6}$.
- Note that this definition is only valid when all outcomes are equally likely.
- Therefore, is it useful to calculate probabilities for the outcome of a fair die $\{1, 2, \cdots, 6\}$, but not for the outcome of a single specific election $\{\text{Candidate A Wins}, \text{Candidate B Wins}\}$.

Challenges with One-Shot Events

- Many events cannot be repeated under identical conditions.
  - Example: A specific election outcome.
- Frequentist and classical definitions struggle with one-shot events.
  - What does it mean to repeat an election?
  - How do we count all possible outcomes for a unique event?
- We need additional assumptions in these cases.
- How we define our population becomes crucial in calculating probabilities.
- This often means we need to make our research question more general.

Classical Probability in the Context of Elections

- **Calculating Probabilities in Electoral Contexts**:
  - Collect data from a large number of past elections within a given country or state.
  - Ensure these elections are comparable (e.g., same country/state, similar conditions).
  - Calculate the probability of a candidate with specific characteristics (e.g., gender, age, race) getting elected.
- **Example**:
  - **Suppose** we have data from 200 **comparable elections**. Define this as the **population**.
  - In 55 of these elections, a woman candidate won.
  - Probability of a woman candidate winning *in our population of elections*:

$$\frac{\text{Number of Favorable Cases}}{\text{Total Number of Cases}} = \frac{110}{200} = 0.55$$

Why it's important to learn probability to understand statistics

Suppose that the Office of Student Life claims that 45% of Stony Brook students are registered to vote in the next presidential election. Then, suppose we take a random sample of 100 students and determine that the proportion registered is:

$$\frac{53}{100} = 0.53 \qquad \text{(That is, 53\%)}$$

Now, we need to ask the question:

- *If the actual population proportion is 0.45, how likely is it that we'd get a sample proportion of 0.53?*
- If the answer is "*fairy likely*," the initial claim is reasonable. If not, we reject it.

Learning about probability allows us to answer such questions.

Definitions

Sample Space. The sample space (or outcome space), denoted S, is the collection of all possible outcomes of a random process (random experiment).

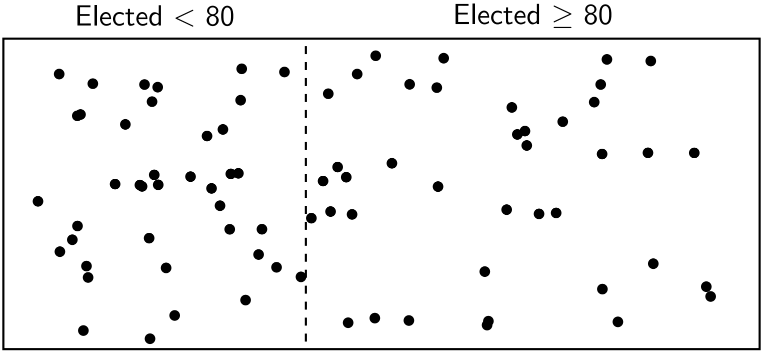
Event. Denoted with capital letters $A, B, C, \ldots$ — is just any subset of the sample space S. For example, $A \subset S$, where "$\subset$" denotes "is a subset of."

Probability of an Event. Let A be an event, such that A includes a subset of outcomes taken from S. Then,

$$\text{Probability of A} = \frac{\text{Total Cases in A}}{\text{Total Cases in S}}$$

Classical Probability in the Context of Elections

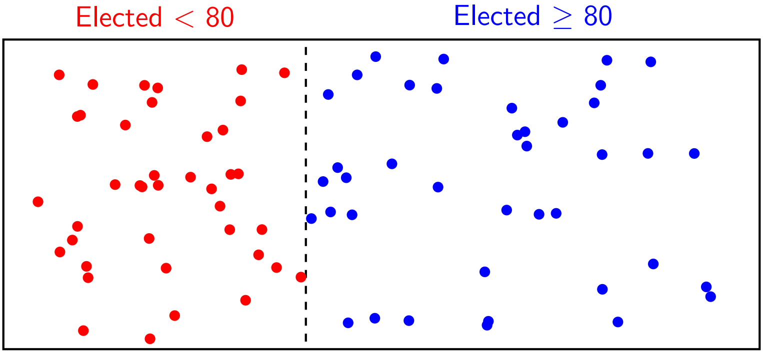
- Population: All elections in the United States in the last 20 years.
- Divided into two parts:
  - Elections in which someone older than 80 was elected.
  - Elections in which someone younger than 80 was elected.
- Dots represent individual cases.



Classical Probability in the Context of Elections

- What is the probability that a candidate younger than 80 years old is elected?

$$P(Elected < 80) = \frac{\#\ \text{(Favorable Cases)}}{\#\ \text{(Total Cases)}} = \frac{\text{Count of Red Dots}}{\text{Total Number of Dots}}$$



Three Axioms of Probability - Introduction

Probability Function

Probability is a real-valued function P that assigns to each event A in a sample space S a number known as the **probability of the event** A, denoted by $P(A)$.

An axiom in mathematics is a foundational statement assumed to be true, serving as a building block for the branch of mathematics in question. In probability and statistics, our theory is based upon three axioms.

Three Axioms of Probability - Detailed

The function P satisfies the following properties (axioms):

1. $P(A) \geq 0$ for any event A. (Nonnegativity)
2. $P(S) = 1$, where S is the sample space. (Certainty)
3. For any sequence, $A_1, A_2, \ldots A_k$, of mutually exclusive events, i.e., $A_j \cap A_i = \emptyset$, ,

   $$P(A_1 \cup A_2 \cup A_3 \cdots \cup A_k) = P(A_1) + P(A_2) + P(A_3) + \cdots + P(A_k)$$

   (Additivity)

These axioms form the basis of probability theory, providing a foundation for all subsequent probability rules.

Probability Distributions for Finite Sample Spaces

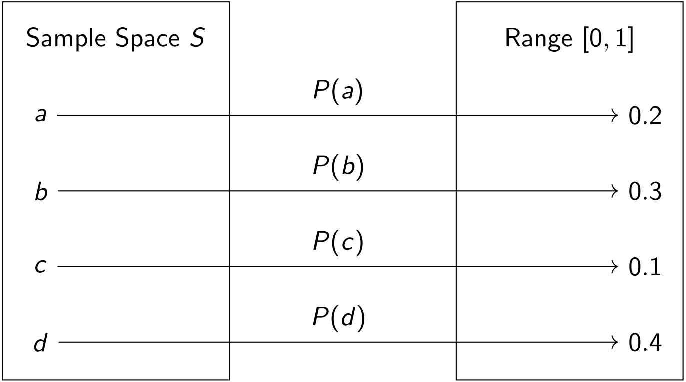
Probability distributions describe how probabilities are assigned across all possible outcomes in a finite sample space.

Definition: A probability distribution assigns a probability $P(a)$ to each outcome a in the sample space such that:

- $P(a) \geq 0$ for each outcome a
- The sum of probabilities for all outcomes is 1, $\sum_{a \in S} P(a) = 1$

The symbol $\in$ is used to construct a statement in set theory notation, and it means "belongs to" or "is an element of." Imagine the set $S = \{a, b, c, d\}$, therefore the statement "$a \in S$" is true because a is an element of S.

Probability Mapping: $P : S \to [0, 1]$



Probability Problem: Number of Heads When Throwing Two Coins

Problem Statement:

- We toss **two fair coins**.
- We define an **event** as a specific **number of heads** observed.
- Our goal is to determine the probability distribution of the different realizations of this outcome.

Sample Space: The possible outcomes of flipping two fair coins are:

$$S = \{HH, HT, TH, TT\}$$

where:

- HH: Both coins land on heads.
- HT and TH: One coin lands on heads, the other on tails.
- TT: Both coins land on tails.

Computing the Probability Distribution of X

Step 1: Count the Outcomes for Each Event

- $X = 0$ (No heads): $\{TT\} \rightarrow$ 1 outcome
- $X = 1$ (One head): $\{HT, TH\} \rightarrow$ 2 outcomes
- $X = 2$ (Two heads): $\{HH\} \rightarrow$ 1 outcome

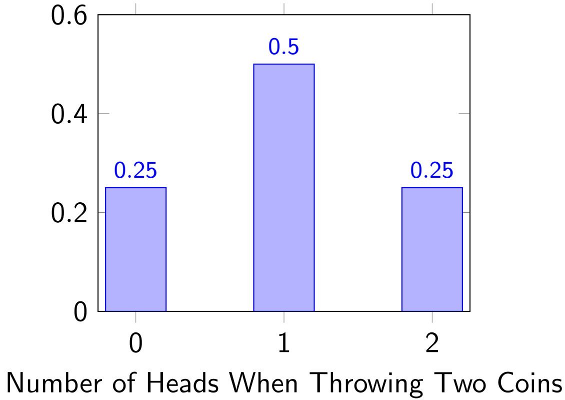
Step 2: Compute Probabilities

X (Number of Heads)	Favorable Outcomes	Probability $P(X)$
0	TT	$1/4 = 0.25$
1	HT, TH	$2/4 = 0.50$
2	HH	$1/4 = 0.25$

Final Answer: The probability distribution of the event is:

$$\text{P(Zero Heads)} = 0.25, \quad \text{P(Just One)} = 0.50, \quad \text{P(Two Heads)} = 0.25$$

Probability Distribution of the Total Number of Heads When Throwing Two Coins



Setting: Probability Distribution of Discrete Outcomes

- We analyze hypothetical (simulated) historical data of country pairs.
- Each combination of variables (Democratic Regimes, Economic Interdependence, War) for a pair of two countries is an outcome.
- We aim to calculate the probability distribution of these discrete outcomes.
- Variables:
  - **Both Countries Are Democratic:** Yes/No (No = if at least one county of the pair is not a democracy)
  - **Economic Interdependence:** High/Low (Do the two countries share strong economic ties and transactions?)
  - **War:** Yes/No (Have the two countries ever been in a war against each other?)

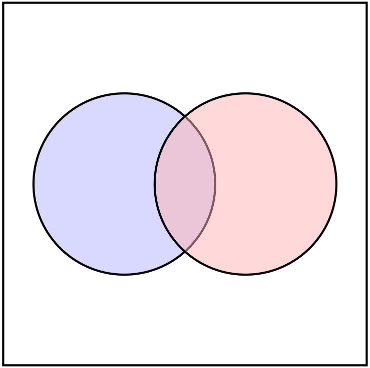
Summary of Dyadic (pairwise) Interactions

Both Democratic	Economic Interdependence	War	Count	Prob.
No	Low	Yes	23	0.12
No	Low	No	82	0.43
No	High	Yes	4	0.02
No	High	No	26	0.14
Yes	Low	Yes	5	0.03
Yes	Low	No	35	0.18
Yes	High	No	15	0.08
Yes	High	Yes	0	0.00

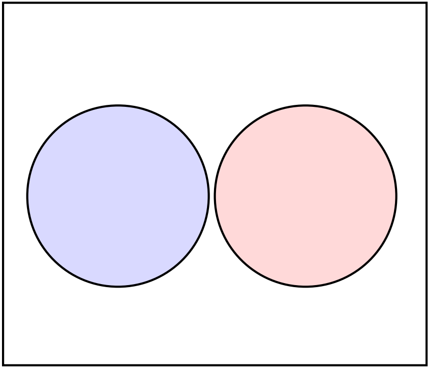
Table: Probability distribution of randomly selecting one country pair combination. Total country pairs = 190.

Venn Diagrams within the Universal Set

Non-empty Intersection



No Intersection



Probability of the Intersection of Two Events

Probability of the Intersection of Two Events $(Def.)$

Following the classical definition, the probability that the event A happens AND the event B happens is given by:

$$P(A \cap B) = \frac{\#(\text{Both } A \text{ and } B \text{ happen})}{\#(\text{Total Cases in } S)}$$

Probability of the Union of Two Events

Probability of the Union of Two Events (*Def.*)

Following the classical definition, the probability that the event A happens OR the event B happens is given by:

$$P(A \cup B) = \frac{\#(A \text{ and not } B) + \#(B \text{ and not } A) + \#(\text{Both } A \text{ and } B)}{\#(\text{Total Cases in } S)}$$

Probability Rules

Addition Rule for Disjoint Events

The addition rule for disjoint (mutually exclusive) events states that if two events A and B cannot occur at the same time, then the probability of A or B occurring is the sum of their probabilities.

$$P(A \cup B) = P(A) + P(B)$$

Example:

- Let A be "At least one country is not a democracy AND They have low economic dependence AND The country pair has never been been at war".
- Let B "Both countries are democratic AND they have low economic dependence AND they have never been at war."
- $P(A \cup B) = P(A) + P(B) = 0.43 + 0.18 = 0.61$

General Addition Rule for Probabilities

The general addition rule is used to find the probability that either of two events occurs.

Formula

If A and B are any two events, then the probability that A or B occurs is given by:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Explanation:

- $P(A \cup B)$ is the probability of either A or B occurring.
- $P(A) + P(B)$ adds the probabilities of A and B occurring.
- Subtracting $P(A \cap B)$ corrects for the double counting of the intersection, where both A and B occur.

Venn Diagrams: Example.

Hypothetical **Dataset:** Comparative sample of countries.
Variables: 1) Regime type (Democratic vs. Not Democratic), 2) Level of Economic Inequality (High vs. Low).

	Inequality		**Row Totals**
Regime	**High**	**Low**	**Total**
Democratic	0.18	0.42	0.60
Not Democ.	0.24	0.16	0.40
Column Totals	0.42	0.58	1.00

Table: Joint Probability Distribution

Question: What is the probability that a randomly selected country is either Democratic or has Low Inequality?

Example

Example:

Let A be "a country is a democracy" and B be "a country has low inequality." Consider a scenario where:

- $P(A) = 0.60$ (Probability of being a democracy),
- $P(B) = 0.58$ (Probability of having low inequality),
- $P(A \cap B) = 0.42$ (Probability of being both Democratic and having low inequality).

Using the general addition rule:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.60 + 0.58 - 0.42 = 0.76$$

Thus, if we select a random country from the sample, the probability that it is either Democratic or has Low Inequality is 76%.

Introduction to Conditional Probability

Overview:

- Conditional probability is a measure of the probability of an event occurring given that another event has already occurred.
- This type of probability is essential in scenarios where the occurrence of one event affects the likelihood of another.

Concept:

- Instead of considering all possible outcomes, conditional probability focuses only on the outcomes where a specific condition or event has occurred.
- Useful in situations like medical testing, where we might be interested in the probability of a disease given a positive test result.

The Formula for Conditional Probability

Conditional Probability Defined:

- The probability of an event A given that event B has occurred is written as $P(A|B)$.

Formula:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- This formula assumes that $P(B) > 0$, i.e., the condition event B has a non-zero probability of occurring.
- Note that $P(A|B) \neq P(B|A)$.

General Multiplication Rule

- Note that this gives a general formula for $P(A \cap B)$

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

Example: Using Conditional Probability

Problem Setup:

- We have a bag with 6 red and 4 blue marbles.
- We draw two marbles **without replacement**.
- We want the probability of drawing a blue marble first, then a red marble.

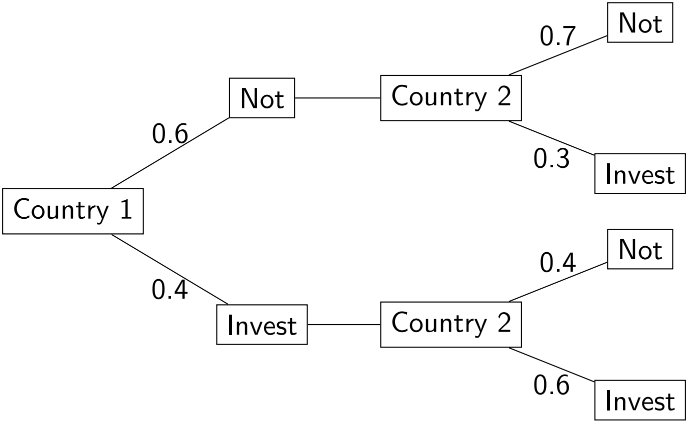
Calculation:

- Probability of drawing a blue marble first: $\frac{4}{10}$.
- Probability of then drawing a red marble (given blue was drawn): $\frac{6}{9}$.
- Combined probability:

$$P(\text{Blue first, then Red}) = \frac{4}{10} \times \frac{6}{9} = \frac{4}{15}.$$

Example: Probability Tree for Nuclear Weapon Investment

We are studying two countries: Country 1 and Country 2. Each country decides whether to invest in nuclear weapons or not.
Probability Tree:



Conditional Probability

What's the probability of the event "Country 2 invests in nuclear weapons (Invest_2), but Country 1 does not (Not_1)"
We can use the following rule: $P(A \cap B) = P(A) \cdot P(B|A)$

Calculation:

- According to the respective branch of the tree: $P(\text{Invest}_2|\text{Not}_1) = 0.3$
- Also, $P(\text{Not}_1) = 0.6$.
- Thus, $P(\text{Invest}_2 \cap \text{Not}_1) = 0.6 \times 0.3 = 0.18$.

Independent Events - Definitions and Properties

Independent Events:

- Events A and B are independent if the occurrence of one does not affect the probability of the occurrence of the other.
- Two events are independent if either:

$$P(B|A) = P(B) \quad \text{(provided that } P(A) > 0\text{) or}$$

$$P(A|B) = P(A) \quad \text{(provided that } P(B) > 0\text{).}$$

- Now, since independence tells us that $P(B|A) = P(B)$, we can substitute $P(B)$ in for $P(B|A)$ in the formula given to us by the multiplication rule:

$$P(A \cap B) = P(A) \times P(B|A) = P(A) \times P(B)$$

Independent Events - *Alternative* Definition

- Given the previous result, events A and B are **independent events** if and only if:

$$P(A \cap B) = P(A) \times P(B)$$

- Otherwise, A and B are called **dependent events**.
- Recall that the "if and only if" in the definition means that the if-then statement works in both directions, in ther words:
  1. If events A and B are independent, then $P(A \cap B) = P(A) \times P(B)$.
  2. If $P(A \cap B) = P(A) \times P(B)$, then events A and B are independent.

Definitions of $P(A \cap B)$ for Independent and Dependent Events

Independent Events

Definition

If events A and B are independent, then:

$$P(A \cap B) = P(A) \times P(B)$$

Dependent Events

Definition

If events A and B are not independent, then:

$$P(A \cap B) = P(A) \times P(B|A)$$

or equivalently:

$$P(A \cap B) = P(B) \times P(A|B)$$

Probability of the Complement

Derivation

Consider an event A and its complement A^c. Since A and A^c are mutually exclusive and exhaustive, the addition rule gives:

$$P(A \cup A^c) = P(A) + P(A^c) - P(A \cap A^c)$$

Because A and A^c are mutually exclusive, $P(A \cap A^c) = 0$. Also, $A \cup A^c$ covers the entire sample space, so $P(A \cup A^c) = P(S) = 1$. Thus:

$$1 = P(A) + P(A^c)$$

Consequently:

$$P(A^c) = 1 - P(A)$$

Example: Law School Applications

Suppose a student estimates a 15% chance of being accepted by any given law school. Assuming acceptance decisions are independent, how many schools should they apply to have a probability of getting accepted to at least one school be higher than 80%?

Let A be the event of being accepted to one particular school, so:

$$P(A) = 0.15, \quad P(A^c) = 0.85.$$

If the student applies to n schools (assuming independent decisions), the probability of being rejected by all schools is:

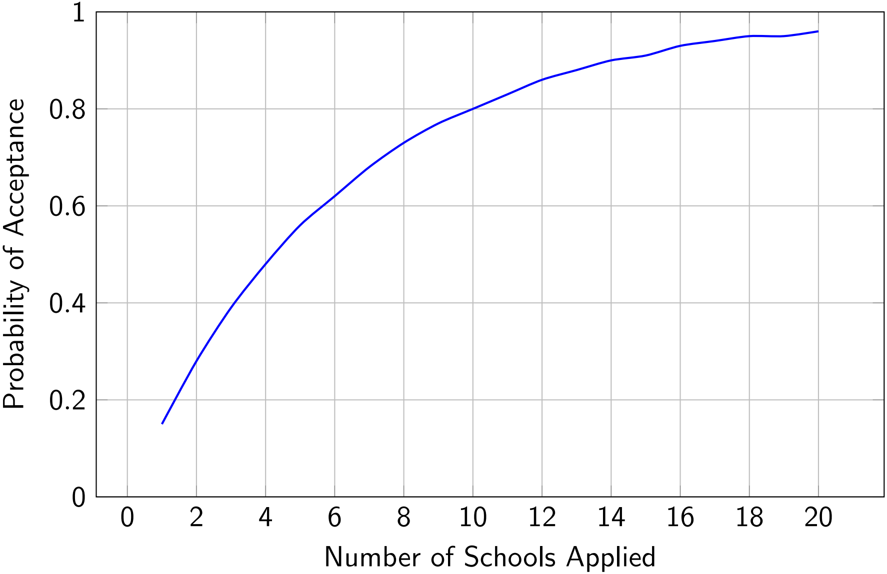
$$P(\text{no acceptances}) = (0.85) \times (0.85) \cdots \times (0.85) = (0.85)^n.$$

Therefore, the probability of at least one acceptance is:

$$P(\text{at least one acceptance}) = 1 - (0.85)^n.$$

We want this probability to be at least 80%. If try for different n we see that 10 is the minimum number of applications such that $P(\text{at least one}) \geq 0.80$

Probability of Acceptance to At Least One School



Law of Total Probability - Intuition

Intuition:

- The Law of Total Probability helps us calculate the probability of an event by considering all possible ways that event can occur.
- It involves breaking down the event into mutually exclusive cases by conditioning on a different event.
- Consider the following example. What's the probability of randomly selecting one Senator and that they are woman?

Law of Total Probability - Example

Define the events B and A. We can split the event B into two cases 1) B and A happen, and 2) B and A^c (not A) happens.

Consider the following example. Assume N_W current senators are woman. Define O as the event "a senator is older than 80 years." And define the event $W \cap O$ as the event "A senator is woman and older than 80 years." Hence, it must be the case that:

- $N_W = N_{W \cap O} + N_{W \cap O^c}$
- Total Woman Senators = # (Woman & Older than 80) + # (Woman & Younger than 80).
- Assume that N is the total number of senators. Then,
- $\frac{N_W}{N} = \frac{N_{W \cap O}}{N} + \frac{N_{W \cap O^c}}{N}$
- $P(W) = P(W \cap O) + P(W \cap O^c)$.

Law of Total Probability - Intuition

Intuition:

- The Law of Total Probability calculates the probability of an event by breaking it down into the mutually exclusive cases the event can happen jointly with another.

Formula:

$$P(B) = P(B \cap A) + P(B \cap A^c)$$

Alternatively, applying the general multiplication rule:

$$P(B) = P(B|A) \cdot P(A) + P(B|A^c) \cdot P(A^c)$$

Explanation:

- $P(B|A)$ is the probability of B given A.
- $P(A)$ is the probability of A.
- $P(A^c)$ is the complement of A.

Law of Total Probability - Example

Example:

- Suppose we have two bags of marbles:
  - Bag One: 70% red marbles, 30% blue marbles.
  - Bag Two: 40% red marbles, 60% blue marbles.
- We randomly choose a bag, with a 50% chance of picking either. What's the probability of drawing a blue marble (B)?

Calculation:

- Let A be the event of choosing Bag One, and A^c be the event of choosing Bag Two.
- $P(B|A) = 0.3$ (Probability of blue marble from Bag One)
- $P(B|A^c) = 0.6$ (Probability of blue marble from Bag Two)
- $P(A) = 0.5$, $P(A^c) = 0.5$

$$P(B) = P(B|A) \cdot P(A) + P(B|A^c) \cdot P(A^c)$$
$$P(B) = 0.3 \cdot 0.5 + 0.6 \cdot 0.5 = 0.15 + 0.3 = 0.45$$

Bayes' Theorem - Intuition and Example

Intuition:

- Bayes' Theorem helps us update our beliefs (probabilities) based on new evidence.
- **Prior:** The initial probability before seeing the new evidence.
- **Posterior:** The updated probability after considering the new evidence.

Bayes' Theorem - Intuition and Example

Example:

- Suppose 2% of politicians are corrupt ($\Pr(C) = 0.02$).
- If a politician is corrupt, there is a 95% chance they are mentioned in the Panama Papers ($\Pr(PP|C) = 0.95$).
- If a politician is not corrupt, there is a 10% chance they are mentioned in the Panama Papers ($\Pr(PP|\neg C) = 0.1$).

Problem: Suppose it is found that a politician was mentioned in the Panama papers. What is the probability that they are corrupt? (i.e., compute $\Pr(C|PP)$).

Bayes' Theorem - General Formula and Law of Total Probability

Bayes' Theorem:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|\neg A) \cdot P(\neg A)}$$

Terms in Bayesian Nomenclature:

- $P(A|B)$ is the posterior probability of A given B.
- $P(B|A)$ is the likelihood of B given A.
- $P(A)$ is the prior probability of A.
- $P(B)$ is the marginal probability of B, computed using the Law of Total Probability.

Panama Papers and Corruption - Bayes Theorem

Calculation:

$$Pr(P) = Pr(P|C) \cdot Pr(C) + Pr(P|\neg C) \cdot Pr(\neg C)$$

$$Pr(P) = (0.95 \times 0.02) + (0.1 \times 0.98) = 0.019 + 0.098 = 0.117$$

$$Pr(C|P) = \frac{P(PP|C) \times P(C)}{P(PP)} = \frac{0.95 \times 0.02}{0.117} \approx 0.162$$

- The probability that a politician is corrupt, given a mention in the Panama Papers, is approximately 16.24%.